Amendments to the Specification:

Please replace paragraph [0005] with the following replacement paragraph:

[0005] The so-called SAGE (Serial Analysis of Gene Expression) method is known as an efficient method of obtaining partial information on the base sequences in mRNAs (Velculescu V. E. et at., Science 270, 484-487 (1995)). According to this method, DNA concatemers are formed by ligating multiple short DNA fragments (initially about 10 bp) containing information on the base sequences at the 3' end of multiple mRNAs, and the base sequences in these DNA concatemers are determined. This is a method for obtaining partial information on the base sequences at the 3' end of multiple mRNAs. When only a short base sequence close to the 3' end is available but the mRNAs itself is already known, the SAGE method can often identify a specific mRNA or gene, although the available base sequence is often as short as about 10 bp. Recently, an improved version of SAGE, the so-called LongSAGE, has been published. This method allows for the cloning of longer SAGE tags (Saha S. et al., Nat. Biotechnol. 20, 508-12 (2002), U.S. patent publication Nos. 20030008290 and 20030049653). The SAGE method is currently in wide use as an important method for analyzing genes expressed in specific cells, tissues or organisms, and SAGE tags are available for reference in the public domain, e.g. under ~~http://~~cgap.nci.nih.gov/SAGE.

Please replace paragraphs [0020] through [0022] with the following replacement paragraphs:

[0020] FIG. 4 shows examples for the ligation of the first linker for the cloning of 5' end specific tags are presented. The examples specify the linkers used according to the protocols described in Examples 1 to 3. Exemplary elongation strands are set forth in the figure and in SEQ ID NOs:94, 96, 98, and 100. Exemplary first strand cDNA molecules are also provided in the figure and SEQ ID NOs:95, 97, 99, and 101.

[0021] FIG. 5 shows examples for the ligation of the second linker for the cloning of 5' end specific tags are presented. The examples specify the linkers used according to the protocols described in Examples 1 to 3. Exemplary elongation strands are set forth in the figure and in

SEQ ID NOs:102, 104, 106, and 108. Exemplary first strand cDNA molecules are also provided in the figure and SEQ ID NOs:103, 105, 107, and 109.

[0022] FIG. 6 shows examples for illustrating the structure of a dimer of 5' end tags prepared in accordance with Examples 1 to 3. Note that in the case of concatemers prepared according to Example 1 different linker sites can be found as XmaJI and XbaI create the same overhangs after digestion, which can be recombined. One example for such a concatemer is given in the figure. Exemplary DNA fragments for the preparation of concatemers are provided in the figure and set forth in SEQ ID NOs:110, 111, 112, 113, 114, 115, 116, 117, 118, and 119.

Please replace paragraph [0062] with the following replacement paragraph:

[0062] Sequences from specific 5' end tags obtained from concatemers in the aforementioned form can be analyzed for their identity by standard software solutions to perform sequence alignments like NCBI BLAST (~~http://~~www.ncbi.nlm.nih.gov/BLAST/), FASTA, available in the Genetics Computer Group (GCG) package from Accelrys Inc. (~~http://~~www.accelrys.com/-), or alike. Such software solutions allow for an alignment of 5' end specific sequence tags among one another to identify unique or non-redundant tags for clustering and further use in database searches. All such non-redundant sequence tags can then be individually counted and further analyzed for the contribution of each non-redundant tag to the total number of all tags obtained from the same sample. The contribution of an individual tag to the total number of all tags should allow for a quantification of the transcripts within a plurality of mRNAs or a cDNA library. The results obtained in such a way on individual samples can be further compared with similar data obtained from other samples to compare their expression patterns against each other. Thus the invention allows for the expression profiling of individual transcripts within one or more samples and the establishment of a reference database.

Please replace paragraph [0064] with the following replacement paragraph:

[0064] Specific 5' end sequence tags which could be mapped to genomic sequences allow for the identification of regulatory sequences (Suzuki Y et al. EMBO Rep. 2001 May; 2(5):388-93 and Suzuki Y et al. Genome Res. 2001 May; 11(5):677-84). In a gene the DNA upstream of the 5'

end of transcribed regions usually encompasses most of the regulatory elements which are used in the control of gene expression. These regulatory sequences can be further analyzed for their functionality by searches in databases which hold information on binding sites for transcription factors. Publicly available databases on transcription factor binding sites and for promoter analysis including Transcription Regulatory Region Database (TRRD) (~~http://~~wwwmgs.bionet.nsc.ru/mgs/dbases/trrd4/), TRANSFAC (~~http://~~transfac.gbf.de/TRANSFAC/), TFSEARCH (~~http://~~www.cbrc.jp/research-/db/TFSEARCH.html), and PromoterInspector provide by Genomatix Software (~~http://~~www.genomatix.de/) provide resources for computational analysis of promoter regions.

Please replace paragraph [0178] with the following replacement paragraph:

[0178] Alternatively or additionally, random primer (dN_6- dN_9), where N is any nucleotide mRNA, recommended 2.5 to 25 μg or alternatively, total RNA, 5-50 μg, and where V refers to a nucleotide selected from the group consisting of A, C, or G.

Please add this sequence identifier following the nucleotide sequence set forth in paragraph [0497]:

[0497] (SEQ ID NO:78)

Please replace paragraph [0499] with the following replacement paragraph:

[0499] Linkers found using "cross~match(excerpts from output):
linker1
TCTAGGTCCGACG 86-98 13-1 C (SEQ ID NO: 31)
linker2
TCTAGGTCCGACG 118-130 13-1 C (SEQ ID NO:79)
linker3
CCTAGGTCCGACG 151-163 13-1 C (SEQ ID NO: 32)
linker4
CCTAGGTCCGACG 158-170 1-13 (SEQ ID NO:80)
linker5
TCTAGGTCCGACG 190-202 1-13 (SEQ ID NO:81)
linker6
CCTAGGTCCGACG 249-261 13-1 C (SEQ ID NO:82)
linker7

CCTAGGTCCGACG 256-268 1-13 (SEQ ID NO:83)
linker8
TCTAGGTCCGACG 288-300 1-13 (SEQ ID NO:84)
linker9
CCTAGGTCCGACG 347-359 13-1 C (SEQ ID NO:85)
linker10
CCTAGGTCCGACG 354-366 1-13 (SEQ ID NO:86)

Please add this sequence identifier following the nucleotide sequence set forth in paragraph [0504]:

[0504] (SEQ ID NO:87)

Please add this sequence identifier following the nucleotide sequence set forth in paragraph [0509]:

[0509] (SEQ ID NO:88)

Please add this sequence identifier following the nucleotide sequence set forth in paragraph [0523]:

[0523] (SEQ ID NO:89)

Please add this sequence identifier following the nucleotide sequence set forth in paragraph [0540]:

[0540] (SEQ ID NO:90)

Please replace paragraph [0549] with the following replacement paragraph:

[0549] 5' end specific sequence tags can be analyzed for their identity by standard software solutions to perform sequence alignments like NCBI BLAST (~~http://~~wvmw.ncbi.nlm.nih.gov/BLAST/), FASTA, available in the Genetics Computer Group (GCG) package from Accelrys Inc. (~~http://~~www.accelrys.com/) or alike. Such software solutions allow for an alignment of 5' end specific sequence tags among one another to identify unique or non-redundant tags, which can be further used in Database searches and building a 5'-end sequence database.

Please add this sequence identifier following the nucleotide sequences set forth in paragraph [0571]:

[0571] (SEQ ID NO:91)

Please add this sequence identifier following the nucleotide sequences set forth in paragraph [0576]:

[0576] (SEQ ID NO:91)

Please add this sequence identifier following the nucleotide sequences set forth in paragraph [0581]:

[0581] (SEQ ID NO:91)

Please add this sequence identifier following the nucleotide sequences set forth in paragraph [0657]:

[0657] (SEQ ID NO:92)

Please add this sequence identifier following the nucleotide sequences set forth in paragraph [0668]:

[0668] (SEQ ID NO:92)

Please add this sequence identifier following the nucleotide sequences set forth in paragraph [0673]:

[0673] (SEQ ID NO:92)

Please add this sequence identifier following the nucleotide sequences set forth in paragraph [0678]:

[0678] (SEQ ID NO:93)

Please add this sequence identifier following the nucleotide sequences set forth in paragraph [0689]:

[0689] (SEQ ID NO:92)

Please replace paragraph [0690] with the following replacement paragraph:

[0690] 5' end specific sequence tags obtained as describe in this Example can be used to identify transcribed regions within genomes for which partial or entire sequences were obtained. Such a search can be performed using standard software solutions like NCBI BLAST (~~http://~~www.ncbi.nlm.ni- h.gov/BLAST/) to align the 5' end specific sequence tags to genomic sequences. In the case of large genomes like those from human, rat or mouse it may be necessary to extend the initial sequence information obtained from concatemers. The use of extended sequences allows for a more precise identification of actively transcribed regions in the genome.

Please replace paragraph [0693] with the following replacement paragraph:

[0693] 5' end sequence tags obtained from the same plurality of mRNAs in a sample or nucleic acid fragments within the same cDNA library can be analyzed by a standard software solution like. NCBI BLAST (~~http://~~www.ncbi.nlm.nih.gov/BLAST/) to identify non-redundant sequence tags as describe in Example 5. All such non-redundant sequence tags can then be individually counted and further analyzed for the contribution of each non-redundant tag to the total number of all tags obtained from the same sample. The contribution of an individual tag to the total number of all tags should allow for a quantification of the transcripts in a plurality of mRNAs in the sample or a cDNA library. The results obtained in such a way on individual samples can be further compared with similar data obtained from other samples to compare their expression patterns.

Please replace paragraphs [0695] through [0698] with the following replacement paragraphs:

[0695] Transcription Regulatory Region Database (TRRD)
(~~http://~~wwwmgs.bionet.nsc.ru/mgs/dbases/trrd4/)

[0696] TRANSFAC (~~http://~~transfac.gbf.de/TRANSFAC/)

[0697] TFSEARCH (~~http://~~www.cbrc.jp/research/db/TFSEARCH.html)

[0698] PromoterInspector provide by Genomatix Software (~~http://~~www.genomatix.de/)